MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 3, 2016/2017

DPS5018 – INTRODUCTION TO PROBABILITY AND STATISTICS

(A95-96, A100-101, B231-236, B240-245)

02 JUNE 2017

03.00 pm – 05.00 pm

(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of **4** pages excluding cover page and appendix.

2. Attempt **ALL FOUR (4)** questions. The marks distribution for each question is given.

3. Write your answers in the answer booklet provided.

4. Key formulae are given in the Appendix.

Question 1

a) Fill in the empty spaces below by using the words given in the box. (5 marks)

i. _______________consists of methods for drawing & measuring the reliability of conclusions about a population based on information obtained from a sample of population.

ii. It is found that, on average, 700 visitors love to visit Melaka during public holiday. The term "visitors" can be referred as __________ .

iii. _______________ is one of the sampling methods which the number of population are widely scattered in a geographical area or location.

iv. _________________ is a sample obtained by simple random sampling.

v. _____________________ occurs when member of the population can be selected more than once.

population	Simple random sample	Inferential statistics	sample
Cluster sampling	Sampling with replacement	Sampling without replacement	Descriptive statistics

b) The following shows the brand of smartphones which have high number of users.

Samsung (S)	Iphone (I)	Lenovo (L)	Vivo (V)	Oppo (O)

List the 10 possible outcomes **(without replacement)** if a customer wants to select 3 out of 5 smartphones. [Hint: Use the abbreviation; S, I, L, V, O] (5 marks)

c) EhAsia & Co. decided to provide best staff award to their eligible staff for each department. Table below shows the number of staffs for each department.

Department	Number of staff
Business Development	260
Sales & Commercialization	140
Human Resources	220
Issues & Infrastructure	200
Innovation & Technology	180

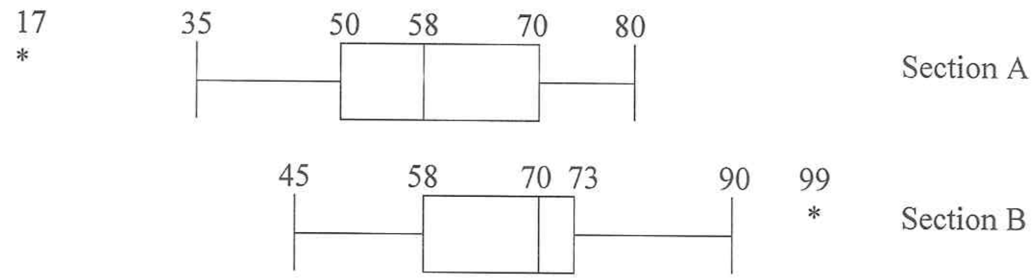
By using **stratified sampling procedure**, obtain the number of staff from each department if it is provided to select 50 samples. (5 marks)

[TOTAL 15 MARKS]

Continued...

Question 2

a) The following boxplot represents the distribution of Statistics examination marks between Section A and Section B.

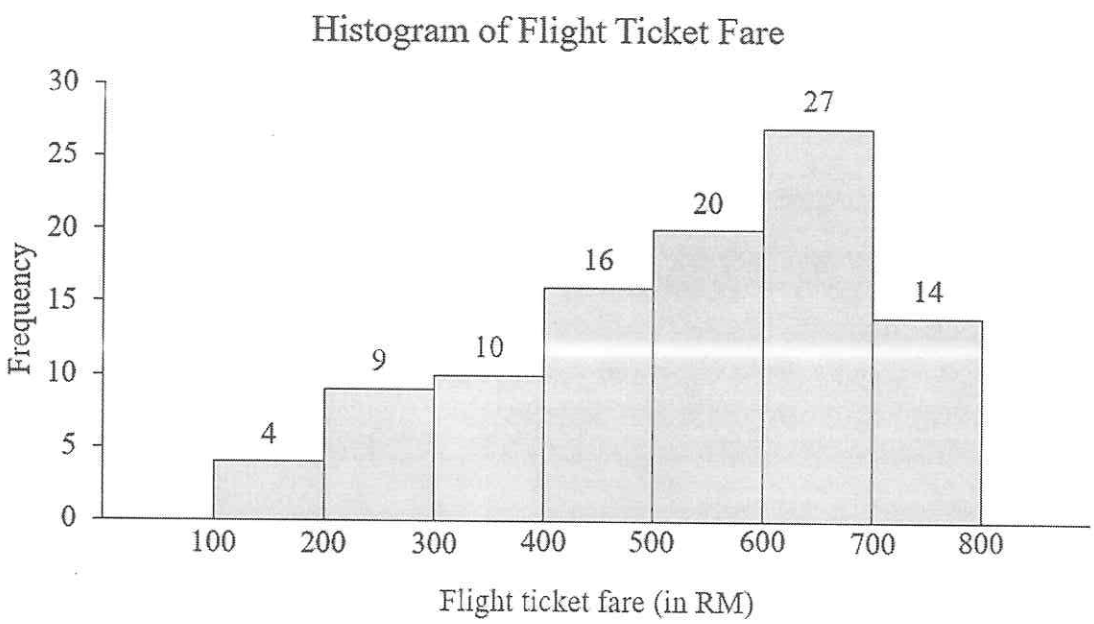


i. According to the diagram above, reconstruct and complete the following table. (6 marks)

	Section A	Section B
Minimum value		
Quartile 1, Q_1		
Quartile 2, Q_2		
Quartile 3, Q_3		
Maximum value		
Outlier		

ii. Calculate the range, lower limit and upper limit for each section. (6 marks)

b) Histogram below shows the respond of a survey to 100 passengers at LCCT on their flight ticket fares (in RM) preference.



Continued...

i. Identify the type of variable in the above situation. (1 mark)

ii. Determine the shape of the distribution. (1 mark)

iii. Based on the histogram, construct a grouped – data – table where the first class starts at 100<200. (6 marks)

iv. From table in part (iii), calculate the variance. Give your answer in three decimal places. (5 marks)

[TOTAL 25 MARKS]

Question 3

a) The following probability distribution recorded the number of persons living in a randomly selected occupied apartment.

Number of persons, x	1	2	3	4	5	6	7
$P(X = x)$	0.16	0.33	0.27	0.06	0.15	0.02	0.01

Find the

i. probability that at least five persons living in a randomly selected occupied apartment. (3 marks)

ii. expected value of X. (2 marks)

iii. standard deviation of X. (3 marks)

b) A lab network consists of 20 computers was attacked by a computer virus. This virus enters each computer with probability 0.6. Find the probability that

i. exactly five computers were attacked by the computer virus. (3 marks)

ii. at most ten computers were attacked by the computer virus. (3 marks)

c) Vehicles pass through a junction on a busy road at an average of 5 vehicles per minute. By assuming the number of vehicles follows a Poisson distribution, find the

i. standard deviation of the distribution. (1 mark)

ii. probability that at least three vehicles will pass through in 1-minute period. (3 marks)

iii. probability that exactly ten vehicles will pass through in 2-minutes period. (3 marks)

d) The time required for a student to complete an examination is normally distributed, with mean of 80 minutes and a standard deviation of 10 minutes. What is the probability a student will complete the exam in less than 75 minutes? (4 marks)

e) Professor Sara recorded the result of Midterm Test 1 and found that the score is normally distributed with a mean of 72 and a standard deviation of 12. If only 15% of her students are able to score grade A, find the minimum marks that a student must obtain in order to get A grade. (5 marks)

[TOTAL 30 MARKS]

Continued...

Question 4

a) Suppose a medical assistant wants to study a relationship between the time taken (in minutes) by ambulance to reach the accident scene and the distance from the hospital (in km). The data of ten emergency calls are given as follows.

Distance (km), x	3.5	2.6	5	3.3	4.2	6.7	1.6	6	1.1	2.1
Time Taken (minutes), y	9	6	11	8	10	14	4	16	3	5

i. Determine the dependent and independent variables. (1 mark)

ii. Given $\sum y^2 = 904, \sum x^2 = 161.61, \sum xy = 380.5$. Compute the value of SS_{xx}, SS_{yy} and SS_{xy}. (3 marks)

iii. Calculate the correlation coefficient, r and interpret the answer. (3 marks)

iv. Compute the regression equation $\hat{y} = a + bx$. (4 marks)

v. Predict the time taken by the ambulance to reach the accident scene if the distance is 8.5km.

(2 marks)

b) A manufacturing process produces ball bearings with diameters that have a normal distribution with standard deviation of 0.04 centimeters. The Quality Department claims that the mean of the ball bearing is differ from 0.50 centimeters. Assume that a random sample of 25 ball bearings has a sample mean diameter of 0.51 centimeters.

i. State the null and alternative hypothesis. (2 marks)

ii. Compute the test statistics, Z and the *p-value*. (3 marks)

iii. Test the claim at 10% level of significance. Explain your answer. (2 marks)

c) The following table shows the survey conducted by MMU Alumni on number of graduates who get a job after six months of diploma graduation. It is claimed that men have less secured on job placement compared to women.

	Men, p_1	Women, p_2
Selected Number of Graduates, x	60	117
Total Graduates, n	240	360

i. Identify the null and alternative hypothesis. (2 marks)

ii. Find the value of $\hat{p}_1$ and $\hat{p}_2$. (2 marks)

iii. From answer in part (ii), calculate the value of $\hat{p}_p$ and obtain the test statistics, Z. (3 marks)

iv. Determine the p-value. (2 marks)

v. At 5% significance level, can we conclude the claim is true? (1 mark)

[TOTAL 30 MARKS]

End of Page.

APPENDIX – KEY FORMULA

	Ungrouped Data	Grouped Data
Mean	$\bar{x} = \dfrac{\sum x_i}{n}$ where n : sample size	$\bar{x} = \dfrac{\sum m_i f_i}{\sum f_i}$ where m : class midpoint f : class frequency
Variance	$s^2 = \dfrac{1}{n-1}\left[\sum x_i^2 - \dfrac{\left(\sum x_i\right)^2}{n}\right]$ where n : sample size	$s^2 = \dfrac{1}{\left(\sum f_i\right)-1}\left[\sum m_i^2 f_i - \dfrac{\left(\sum m_i f_i\right)^2}{\sum f_i}\right]$ where m : class midpoint f : class frequency

DISCRETE RANDOM VARIABLE, BINOMIAL DISTRIBUTION, AND POISSON DISTRIBUTION

Mean of a discrete random variable x	Standard deviation of a discrete random variable x
$\mu = \sum xP(X = x)$	$\sigma = \sqrt{\sum x^2 P(X = x) - \mu^2}$

Binomial probability formula	Poisson probability formula
$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$	$P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$

- Key formula for Binomial Probability Distribution:

If $p \le 0.5$

i. $P(X = r) = B(r) - B(r-1)$

ii. $P(X \ge r) = 1 - B(r-1)$

iii. $P(X \le r) = B(r)$

iv. $P(X > r) = 1 - B(r)$

v. $P(X < r) = B(r-1)$

vi. $P(a \le X \le b) = B(b) - B(a-1)$

If $p > 0.5$

i. $P(X = r) = B(n-r) - B(n-r-1)$

ii. $P(X \ge r) = B(n-r)$

iii. $P(X \le r) = 1 - B(n-r-1)$

iv. $P(a \le X \le b) = B(n-b) - B(n-a-1)$

- Key formula for Poisson Probability Distribution:

  If $\mu = \lambda$
  1. $P(X = r) = Poi(r) - Poi(r-1)$
  2. $P(X \geq r) = 1 - Poi(r-1)$
  3. $P(X \leq r) = Poi(r)$
  4. $P(X > r) = 1 - Poi(r)$
  5. $P(X < r) = Poi(r-1)$
  6. $P(a \leq X \leq b) = Poi(b) - Poi(a-1)$

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \dfrac{x-\mu}{\sigma}$

REGRESSION

- For the **least square regression,** $\hat{y} = a + bx$ where $b = \dfrac{SS_{xy}}{SS_{xx}}$ and $a = \overline{y} - b\overline{x}$

- **Correlation coefficient,** $r = \dfrac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$ where

$SS_{xx} = \sum x^2 - \dfrac{\left(\sum x\right)^2}{n}$	$SS_{yy} = \sum y^2 - \dfrac{\left(\sum y\right)^2}{n}$	$SS_{xy} = \sum xy - \dfrac{\left(\sum x\right)\left(\sum y\right)}{n}$

HYPOTHESIS TESTING

Population	Mean	Proportion
1	$Z = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$	$Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$ where $\hat{p} = \dfrac{x}{n}$
2	$Z = \dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$Z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}_p(1-\hat{p}_p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $where \;\; \hat{p}_p = \dfrac{x_1 + x_2}{n_1 + n_2}$